

Lévy area logistic expansion and simulation

BY SIMON J.A. MALHAM AND ANKE WIESE

*Maxwell Institute for Mathematical Sciences and School of Mathematical and
Computer Sciences, Heriot-Watt University, Edinburgh EH14 4AS, UK
(30th June 2011)*

We present a new representation for the Lévy chordal area for a two-dimensional Wiener process conditioned on its endpoints. This is based on an infinite weighted sum of Logistic random variables. We develop a numerical simulation algorithm for the Lévy area based on truncating this series and simulating the tail by a suitable Normal random variable. We show how to improve the efficiency of the algorithm by approximating higher order terms, which are large sums of independent identically distributed Logistic random variables, by two separate methods: using a suitable Normal approximation and, sampling directly from the fixed density function for the logarithm of the product of decimal magnitudes of independent identically distributed Exponential random variables. To implement a strong Milstein numerical integrator for a stochastic differential equation driven by a multi-dimensional Wiener process, we must maintain a local mean-square error of order the cube of the stepsize. To achieve this prescribed accuracy, the latter two Lévy area sampling methods we propose, reduce the number of uniform random variables required to be sampled over each timestep, a measure of the complexity, from reciprocal square-root complexity to logarithmic complexity.

Keywords: Lévy area simulation, logistic expansion

1. Introduction

Fundamental and crucial to high order strong simulation of stochastic differential equations driven by a multi-dimensional Wiener process, is the need to accurately and efficiently simulate the Lévy chordal areas when the diffusion vector fields do not commute. Thusfar this has represented a substantial technical difficulty. In the leading implemented methods, for example Lévy's Normal expansion with a suitable Normal approximation of the tail sum, the number of random variables required to accurately simulate the Lévy chordal area in the Milstein method, scales like the inverse square-root of the stepsize (Wiktorsson 2001). Here we introduce three new simulation methods for the Lévy area based on a new expansion in terms of Logistic random variables. Beyond a given accuracy threshold, all three new simulation methods deliver highly accurate Lévy area samples with orders of magnitude less computational effort. Indeed for two of the methods, the accuracy achievable scales logarithmically with the computational effort.

Consider the problem of strongly simulating a stochastic differential equations driven by two independent Wiener processes W_t^1 and W_t^2 . Assume that the corresponding diffusion vector fields do not commute. To implement an order one

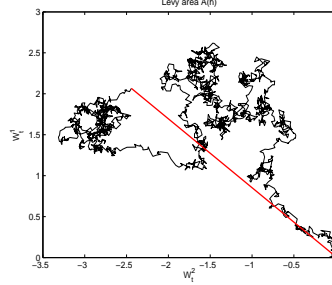


Figure 1. Sample Lévy chordal area—this is the area between the two-dimensional Wiener process shown and the chord joining the endpoints.

Milstein method we must, on each computational timestep of size h , generate two independent $N(0, \sqrt{h})$ sample Wiener increments, $\Delta W^1(h)$ and $\Delta W^2(h)$, and also a sample of the Lévy area defined by

$$A(h) := \frac{1}{2} \int_t^{t+h} \int_t^{\tau_1} dW_{\tau_2}^1 dW_{\tau_1}^2 - dW_{\tau_2}^2 dW_{\tau_1}^1.$$

By Stokes' Theorem this is the area enclosed by the two-dimensional Wiener path and the chord joining the endpoints of the path at t and $t+h$; see Figure 1.

In his seminal 1951 paper, Lévy introduced an equivalent definition of the chordal area as follows. Using the Fourier series representation for the Brownian bridges based on the Wiener processes W_t^1 and W_t^2 given $\Delta W^1(h)$ and $\Delta W^2(h)$, Lévy showed that the chordal area has the series expansion

$$A(h) = \frac{h}{2\pi} \sum_{k=1}^{\infty} \frac{1}{k} \left(U_k(Y_k - \sqrt{\frac{2}{h}} \Delta W^2(h)) - V_k(X_k - \sqrt{\frac{2}{h}} \Delta W^1(h)) \right),$$

where U_k, V_k, X_k, Y_k are independent standard Normal random variables. More generally orthogonal eigenfunction series expansions for Wiener processes are referred to as Karhunen–Loève decompositions after independent work by Karhunen (1947) and Loève (1945). In the same 1951 paper, Lévy proved that the *characteristic function* $\hat{\phi} = \hat{\phi}(\xi)$ corresponding to the probability density function ϕ for the Lévy area $A(h)$, given $\Delta W^1(h)$ and $\Delta W^2(h)$, is

$$\hat{\phi}(\xi) = \frac{\frac{1}{2}h\xi}{\sinh(\frac{1}{2}h\xi)} \exp\left(-\frac{1}{2}a^2\left(\frac{1}{2}h\xi \coth(\frac{1}{2}h\xi) - 1\right)\right),$$

where $a^2 = ((\Delta W^1(h))^2 + (\Delta W^2(h))^2)/h$.

Kloeden, Platen & Wright (1992) suggested truncating the Lévy's series expansion above to include N terms, and using that to generate approximate samples. The mean-square error associated such a truncation scales like h^2/N . However in an important advance, Wiktorsson (2001) proposed simulating the tail sum by a matched Normal random variable $(h/\sqrt{2\pi}) (\pi^2/6 - \sum_{k=1}^N 1/k^2)^{1/2} Z$ where $Z \sim N(0, 1)$. With this correction, the mean-square error scales like h^2/N^2 . In fact Wiktorsson achieves much more than this, deriving an effective sampling method

for the Lévy areas involved in a Milstein approximation of a stochastic differential equation driven by any high-dimensional Wiener process; see Gilling & Shardlow (2007) for a practical implementation. In this paper we will restrict ourselves to a two-dimensional Wiener process.

We measure the effectiveness of any such simulation methods as follows. To successfully implement the Milstein method, we must ensure that the mean-square local truncation error is of order h^3 . Hence over each computational timestep, if we used Wiktorsson's method above, we would need to take $N \geq h^{-1/2}$. We will measure the computational effort associated with such an approximation, by the number of uniform random variables ε that are required to generate such a Lévy area sample. We will call this the *complexity* of the simulation method; also see Rydén & Wiktorsson (2001). Roughly, for Wiktorsson's method, the number of uniform samples we require scales like N . Note, we ignore: (1) the negligible effort associated with converting the uniform samples to standard Normal samples; (2) constant multiplicative factors (there are four standard Normal samples required at each order) and (3) the negligible effort associated with sampling the tail. Hence for Wiktorsson's method, the complexity required to achieve accuracy of order h^3 is $\varepsilon^{-1/2}$. The smaller the complexity, the more effective is the simulation method.

Gaines & Lyons (1994) proposed Marsaglia's rectangle-wedge-tail method for generating Lévy area samples. However, Wiktorsson (2001) observes that "this method is complicated to implement, however, and occasionally requires numerical inversion of the characteristic function". Gaines & Lyons followed this in 1997 with an approximate simulation method based on replacing the Lévy area by its conditional expectation on intervening Brownian path information (also see Lord, Malham & Wiese 2008 and Malham & Wiese 2008). The complexity of this latter method is ε^{-1} .

Rydén & Wiktorsson (2001) developed a method based on an expansion for the Lévy area in Laplace random variables. Observing that the characteristic function $\hat{\phi}$ is the product of the characteristic function of a Logistic random variable and a Poisson mix of Laplace random variables, they derived the representation

$$A(h) = \frac{h}{2\pi} \left(X + \sum_{k=1}^{\infty} \frac{1}{k} \sum_{j=1}^{P_k} Y_{jk} \right),$$

where $X \sim \text{Logistic}(1)$, and for each $k \in \mathbb{N}$ we have: $P_k \sim \text{Poisson}(a^2)$ and for $j = 1, \dots, P_k$, $Y_{jk} \sim \text{Laplace}(1)$; all independent. Note that the expectation of P_k is a^2 , independent of k . In a practical simulation we would truncate this series to include all the terms for $k \leq N$ and simulate the tail sum by the Normal random variable $(h/2\pi) a \sqrt{2/k} Z$ where $Z \sim \mathcal{N}(0, 1)$. This simulation method has complexity $\varepsilon^{-1/2}$.

Some related recent papers of interest are as follows. In a promising paper Stump & Hill (2005) derived a series expansion for the Lévy area density function ϕ based on analytically computing the inverse Fourier transform of the characteristic function $\hat{\phi}$. The terms in the resulting series expansion decay like $1/k^2$ with k as the summation index (see page 407 in Stump & Hill). However they only provide comparisons of the distribution functions and do not implement a sampling method. Levin & Wildon (2008) have developed methods for generating moments of the Lévy area utilizing combinatorial shuffle product structures. Also, assuming Hörmander's

ellipticity condition, Cruzeiro, Malliavin & Thalmaier (2004) have shown how to orthogonally evolve the frame bundle associated with the underlying system to avoid computing the Lévy area entirely.

Let us now summarize the ideas underlying the simulation methods we propose. Our starting point is Rydén & Wiktorsson's expansion for the characteristic function $\hat{\phi}$ of the Lévy area. As already noted, the first factor in the representation for $\hat{\phi}$ above, is the characteristic function associated with a Logistic random variable. We thus focus on the second factor, and for the moment, the expression in the exponent. Using an elementary hyperbolic function identity which permits indefinite iteration, we derive an infinite series expansion for the exponent with terms of the form $(\frac{1}{2}a^2 2^n) (z/2^n)/\sinh(z/2^n)$ where n is the summation index. Applying the exponential function to this series for the exponent generates an infinite product of exponentials of such terms. Now we first observe that characteristic functions of exponential form correspond to a Poisson mix of random variables whose characteristic function is the exponent. Then second we observe that the exponents (the individual terms from the series) are characteristic functions corresponding to scaled Logistic random variables. Hence we can represent the Lévy area, given $\Delta W^1(h)$ and $\Delta W^2(h)$, as the following series of Logistic random variables (see Theorem 2.1 below):

$$A(h) \sim \frac{h}{2\pi} \left(X + \sum_{n=0}^{\infty} \frac{1}{2^n} \sum_{k=1}^{P_n} X_{n,k} \right),$$

where for $n = 0, 1, 2, \dots$, the $P_n \sim \text{Poisson}(\frac{1}{2}a^2 2^n)$ are independent Poisson random variables, and for $n = 0, 1, 2, \dots$ and $k = 1, 2, \dots, P_n$, $X, X_{n,k} \sim \text{Logistic}(1)$ are independent identically distributed standard Logistic random variables (mean zero, variance $\pi^2/3$).

This representation is the basis of the three simulation methods we propose. The first simulation method is a simple truncation of this series representation together with a tail approximation. Note that at each order n , we must on average generate an increasing number, namely $\frac{1}{2}a^2 2^n$ uniform random variables. This means that the complexity of this method is ε^{-1} without, and $\varepsilon^{-1/2}$ with, the tail approximation. Hence, in principle this does not represent an improvement in complexity over the Lévy or Rydén–Wiktorsson methods, though in practice the computational effort required to achieve a given accuracy is far superior to the Rydén–Wiktorsson method, and beyond a high accuracy threshold, also superior to the Lévy method. See Figure 2 where we provide an explicit accuracy versus computational effort comparison. The second method we propose, the *Normal approximation*, recognises that at each order we must sum an increasing number of Logistic random variables. Invoking the Central Limit Theorem, we replace sums of Logistic random variables over a threshold number by the appropriate Normal approximation. This of course has a dramatic effect on the computational effort required, without in practice seemingly compromising the accuracy of the simulation method—see Figure 2. The third method we propose, the *Exponential product approximation*, recognises that a sum of P independent identically distributed standard Logistic random variables can be represented by the difference of the logarithms of products of P independent standard Exponential random variables. We derive an explicit asymptotic form for the density function for large P for the logarithm of the product of P Exponential

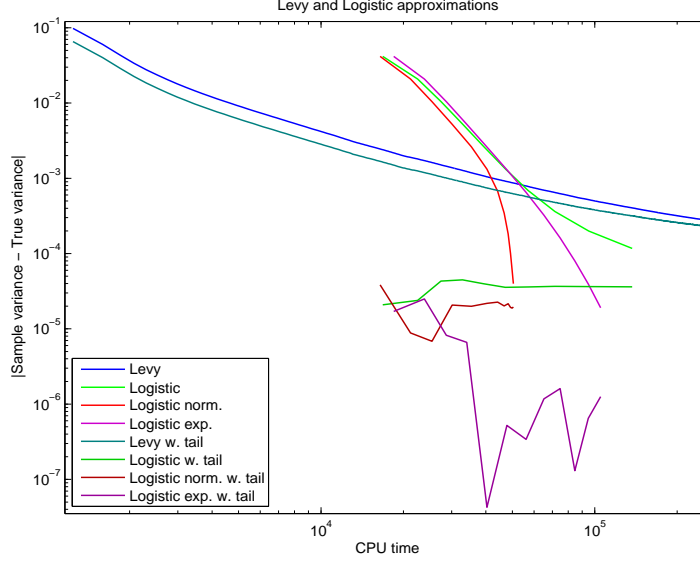


Figure 2. The absolute value of the difference between the sample variance and the true variance $(1 + a^2)h^2/12$ versus the CPU time required to compute the simulation with 5×10^8 independent samples. Here we fixed $h = 1$, so with random Wiener increments $\Delta W^1(h)$ and $\Delta W^2(h)$ for each sample, we set $a^2 = \mathbb{E}\{(\Delta W^1(h))^2 + (\Delta W^2(h))^2\} = 2$.

random variables, using the method of steepest descents. We utilize this explicit form to tabulate the inverse distribution function for $P = 10^2, 10^3, 10^4$ and 10^5 . Combining samples from these tabulated inverse distribution functions with small sums ($< 10^2$) of Logistic random variables generates the highly effective Exponential product approximation shown in Figure 2.

Our paper is structured as follows. We derive our Logistic expansion for the characteristic function for the Lévy area in §2 and provide an algorithm for its implementation including a tail approximation. Then in §3 and §4 we respectively derive the Normal and Exponential product approximations based on the Logistic expansion. In both sections we discuss the complexity of these two methods. Lastly in §5 we provide further simulation details and conclude.

2. Logistic expansion

We propose a different expansion to that considered by Rydén and Wiktorsson (2001) producing a series representation for $A(h)$ with quite different properties.

Theorem 2.1. *The Lévy area $A(h)$ conditioned on the given Wiener increments $\Delta W^1(h)$ and $\Delta W^2(h)$ is equivalent in distribution to the following series of Logistic random variables:*

$$A(h) \sim \frac{h}{2\pi} \left(X + \sum_{n=0}^{\infty} \frac{1}{2^n} \sum_{k=1}^{P_n} X_{n,k} \right),$$

where for $n = 0, 1, 2, \dots$, the $P_n \sim \text{Poisson}(\frac{1}{2}a^2 2^n)$ are independent Poisson random variables, and for $n = 0, 1, 2, \dots$ and $k = 1, 2, \dots, P_n$, $X, X_{n,k} \sim \text{Logistic}(1)$ are independent identically distributed Logistic random variables. Any $X \sim \text{Logistic}(1)$ is efficiently simulated by $X \sim \log(U/(1-U))$ where $U \sim \text{Unif}([0, 1])$.

Proof. We begin by observing the following elementary identity for hyperbolic functions, $\coth z \equiv \coth z/2 - 1/\sinh z$. Iterated N times, this generates the identity

$$\coth z \equiv \coth z/2^{N+1} - \sum_{n=0}^N \frac{1}{\sinh z/2^n}.$$

Hence we see that

$$\begin{aligned} & \exp\left(-\frac{1}{2}a^2(z \coth z - 1)\right) \\ &= \exp\left(\frac{1}{2}a^2 \sum_{n=0}^N 2^n \cdot \frac{z/2^n}{\sinh z/2^n}\right) \cdot \exp\left(-\frac{1}{2}a^2(z \coth(z/2^{N+1}) - 1)\right) \\ &= \prod_{n=0}^N \exp\left(\frac{1}{2}a^2 2^n \cdot \frac{z/2^n}{\sinh z/2^n}\right) \exp\left(-\frac{1}{2}a^2 2^n\right) \cdot \exp(\mathcal{E}_N(z, a)), \end{aligned}$$

where we set

$$\mathcal{E}_N(z, a) := -\frac{1}{2}a^2(z \coth(z/2^{N+1}) - 2^{N+1}).$$

Note that $\mathcal{E}_N(z, a) \rightarrow 0$ as $N \rightarrow \infty$, for all $x \in \mathbb{R}$. We have thus established that for all $\xi \in \mathbb{R}$, setting $z = \xi h/2$, we have

$$\hat{\phi}_{A(h)}(\xi) = \frac{\xi h/2}{\sinh \xi h/2} \cdot \prod_{n=0}^{\infty} \mathbb{E} \left(\frac{\xi h/2^{n+1}}{\sinh \xi h/2^{n+1}} \right)^{P_n},$$

where $P_n \sim \text{Poisson}(\frac{1}{2}a^2 2^n)$ and \mathbb{E} denotes expectation. Now note that the expression $(\xi h/2^{n+1})/\sinh(\xi h/2^{n+1})$ is the characteristic function corresponding to the random variable $\text{Logistic}(h/2^{n+1}\pi) \sim (h/2^{n+1}\pi) \cdot \text{Logistic}(1)$. The result now follows by interpreting the product of two characteristic functions as the sum of two corresponding independent random variables. \square

Hence our proposed simulation method for the Lévy area is as follows. Truncate the logistic expansion in Theorem 2.1 to include the terms $n = 0, 1, \dots, N$. Hence consider the approximation $A_N(h)$ to the conditioned Lévy area $A(h)$ given by

$$A_N(h) \sim \frac{h}{2\pi} \left(X + \sum_{n=0}^N \frac{1}{2^n} \sum_{k=1}^{P_n} X_{n,k} \right),$$

where for $n = 0, 1, \dots, N$: the $P_n \sim \text{Poisson}(\frac{1}{2}a^2 2^n)$ are independent Poisson random variables, for $k = 1, 2, \dots, P_n$, $X = \log(U/(1-U))$ and $X_{n,k} = \log(U_{n,k}/(1-U_{n,k}))$ with $U, U_{n,k} \sim \text{Unif}([0, 1])$ independent identically distributed uniform random variables, and $Z \sim \mathcal{N}(0, 1)$ is a standard Normal random variable. We now consider how to simulate the tail sum.

Theorem 2.2. *The mean-square error of the Logistic expansion approximation $A_N(h)$ is exactly given by*

$$\mathbb{E} |A(h) - A_N(h)|^2 = \frac{a^2}{3 \cdot 2^{N+1}} \left(\frac{h}{2}\right)^2.$$

If we supplement $A_N(h)$ by including the term $(a/\sqrt{3 \cdot 2^{N+1}})(h/2)Z$ to simulate the tail sum, where $Z \sim N(0,1)$, then the mean-square error in the Logistic expansion approximation with tail simulation is

$$\mathbb{E} \left| A(h) - \left(A_N(h) + \frac{a}{\sqrt{3 \cdot 2^{N+1}}} \frac{h}{2} Z \right) \right|^2 \leq \frac{8}{15} \frac{1}{2^{2N+2}} \left(\frac{h}{2}\right)^2.$$

Proof. We directly compute

$$\begin{aligned} \mathbb{E} |A(h) - A_N(h)|^2 &= \mathbb{E} \left(\frac{h}{2\pi} \sum_{n=N+1}^{\infty} \frac{1}{2^n} \sum_{k=1}^{P_n} X_{n,k} \right)^2 \\ &= \left(\frac{h}{2\pi}\right)^2 \sum_{n=N+1}^{\infty} \frac{1}{2^{2n}} \sum_{l=0}^{\infty} \mathbb{P}\{P_n = l\} \cdot \mathbb{E} \left(\sum_{k=1}^l X_{n,k} \right)^2 \\ &= \left(\frac{h}{2\pi}\right)^2 \sum_{n=N+1}^{\infty} \frac{1}{2^{2n}} \sum_{l=0}^{\infty} \mathbb{P}\{P_n = l\} \cdot l \cdot \frac{\pi^2}{3} \\ &= \frac{1}{3} \left(\frac{h}{2}\right)^2 \sum_{n=N+1}^{\infty} \frac{1}{2^{2n}} \cdot \mathbb{E}\{P_n\} \\ &= \frac{1}{3} \left(\frac{h}{2}\right)^2 \sum_{n=N+1}^{\infty} \frac{1}{2^{2n}} \cdot a^2 2^{n-1} \\ &= \frac{a^2}{3} \left(\frac{h}{2}\right)^2 \frac{1}{2^{N+1}}. \end{aligned}$$

In order to obtain the error bound for the Normal tail sum approximation, we note first that the tail sum has class G distribution. Recall that an infinitely divisible random variable has class G distribution, if its characteristic function has the form $\hat{\phi}(\xi) = \exp(-\Psi(\xi^2))$, where $\Psi(0) = 0$, and $(-1)^{n-1} \Psi^{(n)}(\xi) \geq 0$ for all n , see Rydén & Wiktorsson (page 163). Using that

$$A(h) - A_N(h) = \frac{h}{2\pi} \sum_{n=N+1}^{\infty} \frac{1}{2^n} \sum_{k=1}^{P_n} X_{n,k} = \frac{\tilde{h}}{2\pi} \sum_{n=0}^{\infty} \frac{1}{2^n} \sum_{k=1}^{Q_n} X_{n,k},$$

where $\tilde{h} = h/2^{N+1}$, and where Q_n has a Poisson distribution with parameter $\frac{1}{2} \tilde{a}^2 2^n$ and where $\tilde{a}^2 = a^2 2^{N+1}$, it follows that the tail sum $A(h) - A_N(h)$ has the characteristic function (see Lévy 1951)

$$\hat{\phi}_N(\xi) = \exp\left(-\frac{1}{2} \tilde{a}^2 \left(\frac{1}{2} \tilde{h} \xi \coth\left(\frac{1}{2} \tilde{h} \xi\right) - 1\right)\right) = \exp\left(-\frac{1}{2} \tilde{a}^2 \sum_{n=1}^{\infty} \frac{\xi^2}{(n \tilde{h} / 2\pi)^2 + \xi^2}\right).$$

Hence $A(h) - A_N(h)$ has class G distribution—see also Proposition 5 in Rydén & Wiktorsson (2001). We can now proceed as in the proof of Theorem 7 in Rydén & Wiktorsson. The tail sum can be represented as a product of a standard Normal random variable Z and the square root of an independent positive, infinitely divisible variable Y_N , i.e. $A(h) - A_N(h) = Z\sqrt{Y_N}$. If σ_N^2 denotes the variance of $A(h) - A_N(h)$, then the mean-square error when including the Normal tail approximation is given by

$$\begin{aligned} \mathbb{E} |A(h) - A_N(h) - \sigma_N Z|^2 &= \mathbb{E} \{Z^2\} \cdot \mathbb{E} |\sqrt{Y_N} - \sigma_N|^2 \\ &= \mathbb{E} \left(\frac{Y_N - \sigma_N^2}{Y_N + \sigma_N^2} \right)^2 \\ &\leq \frac{1}{\sigma_N^2} \mathbb{E} (Y_N - \sigma_N^2)^2. \end{aligned}$$

Let g denote the Laplace transform of Y_N . Then $g(z) = \hat{\phi}_N(\sqrt{2z})$, and the variance of Y_N is given by $(\log g)''(0)$. If $\ell(z) := -(\sqrt{2z} \coth(\sqrt{2z}) - 1)$ then we see that

$$\log g(z) = \tilde{a}^2 \ell\left(z \frac{\tilde{h}^2}{4}\right) \quad \text{and} \quad (\log g)''(0) = \tilde{a}^2 \left(\frac{\tilde{h}}{2}\right)^4 \ell''(0).$$

Thus we see that

$$\mathbb{E} |A(h) - A_N(h) - \sigma_N Z|^2 \leq \frac{1}{\sigma_N^2} \tilde{a}^2 \left(\frac{\tilde{h}}{2}\right)^4 \ell''(0) = \frac{8}{15} \frac{1}{2^{2N+2}} \left(\frac{h}{2}\right)^2,$$

giving the required result. \square

Note that our expression for the mean-square error of our Logistic approximation $A_N(h)$ is *exact*. Our proposed Logistic approximation to $A(h)$, including simulation of the tail sum is thus $A_N(h) + (a/\sqrt{3 \cdot 2^{N+1}}) (h/2) Z$, where $Z \sim \mathcal{N}(0, 1)$. At each order n in the Logistic approximation we must on average sample $\mathbb{E} \{P_n\} = \frac{1}{2} a^2 2^n$ uniform random variables. Hence the complexity of the Logistic approximation $A_N(h)$ is ε^{-1} , while with tail simulation it is $\varepsilon^{-1/2}$.

3. Normal approximation

We have seen that when we measure the complexity of a simulation method in terms of the number of uniform random variables required to achieve a given accuracy, the Logistic, Laplace (Rydén–Wiktorsson) and Lévy expansion simulation methods scale in the same way. Here we attempt to exploit the form of the Logistic expansion simulation method to significantly improve its complexity competitiveness. For each $n \geq 0$ in the Logistic expansion we must generate and accumulate P_n uniform independent identically distributed random variables where P_n is a Poisson random variable with mean $\mathbb{E} \{P_n\} = \frac{1}{2} a^2 2^n$. Thus on average the number of uniform random variables we must generate and accumulate grows exponentially with n . One possibility to achieve enormous reductions in computational effort for large n , is to proceed as follows. Suppose we generate P_n and realize $P_n \geq P$ for a predetermined $P \gg 1$ to be defined presently. Then in this case we approximate

the Logistic random variable accumulator by a suitable Normal random variable. More precisely, for each $n \geq 0$ for which $P_n \geq P$ we approximate

$$\sum_{k=1}^{P_n} X_{n,k} \approx \sqrt{P_n/3} \pi Z_n,$$

where the $Z_n \sim \mathcal{N}(0, 1)$ are independent standard Normal random variables. Using the Berry–Esseen Theorem we can derive a uniform estimate of the error in the distribution function due to this replacement.

Theorem 3.1. *Let Y_N denote the random variable for some given $P \in \mathbb{N}$:*

$$Y_N(h) := \frac{h}{2\pi} \left(X + \sum_{n=0}^N \frac{1}{2^n} \left(\sum_{k=1}^{P_n} X_{n,k} \cdot 1_{\{P_n < P\}} + \sqrt{P_n/3} \pi Z_n \cdot 1_{\{P_n \geq P\}} \right) \right),$$

and let F_Y denote its distribution function. Let F_A denote the distribution function of $A_N(h)$. Then there exists a constant c such that for all $x \in \mathbb{R}$,

$$|F_A(x) - F_Y(x)| \leq c \frac{N+1}{\sqrt{P}}.$$

Proof. The assertion will follow from the Berry–Esseen Theorem. Indeed, let F_n denote the distribution function of $(1/2^n) \sum_{k=1}^{P_n} X_{n,k} \cdot 1_{\{P_n < P\}} + \sqrt{P_n/3} \cdot \pi \cdot Z_n \cdot 1_{\{P_n \geq P\}}$ and f_n its probability density function, and let G_n denote the distribution function of $(1/2^n) \sum_{k=1}^{P_n} X_{n,k}$ and g_n its probability density function. Thus the distribution functions F_Y and F_A are given as follows (here ‘ \star ’ denotes the usual convolution between probability density functions corresponding to the sum of two random variables):

$$F_Y(x) = \mathbb{P}\{Y_N(h) \leq x\} = \int_{-\infty}^{2\pi x/h} (f_X \star f_1 \star \cdots \star f_N)(y) dy,$$

and

$$F_A(x) = \mathbb{P}\{A_N(h) \leq x\} = \int_{-\infty}^{2\pi x/h} (f_X \star g_1 \star \cdots \star g_N)(y) dy,$$

where f_X denotes the probability density function of the standard Logistic random variable X . The distribution function F_n is given by

$$\begin{aligned} F_n(x) &= \sum_{l=0}^{\infty} \mathbb{P}\{P_n = l\} \cdot \mathbb{P}\left\{ \sum_{k=1}^l X_{n,k} \cdot 1_{\{l < P\}} + \sqrt{l/3} \pi Z_n \cdot 1_{\{l \geq P\}} \leq x 2^n \right\} \\ &= \sum_{l=0}^{P-1} \mathbb{P}\{P_n = l\} \cdot \mathbb{P}\left\{ \sum_{k=1}^l X_{n,k} \leq x 2^n \right\} + \sum_{l=P}^{\infty} \mathbb{P}\{P_n = l\} \cdot \mathbb{P}\left\{ \sqrt{l/3} \pi Z_n \leq x 2^n \right\}, \end{aligned}$$

and the distribution function G_n by

$$G_n(x) = \sum_{l=0}^{\infty} \mathbb{P}\{P_n = l\} \cdot \mathbb{P}\left\{ \sum_{k=1}^l X_{n,k} \leq x \cdot 2^n \right\}.$$

By applying the Central Limit Theorem and the Berry-Esseen Theorem to the sum of logistic random variables $X_{n,k}$ on the right, we know that there exists a constant c , independent of n , such that for all l and for all x

$$\begin{aligned} & |F_n(x) - G_n(x)| \\ &= \left| \sum_{l=P}^{\infty} \mathbb{P}\{P_n = l\} \cdot \left(\mathbb{P}\left\{ \sum_{k=1}^l X_{n,k} \leq x 2^n \right\} - \mathbb{P}\left\{ \sqrt{l/3} \pi Z_n \leq x 2^n \right\} \right) \right| \\ &\leq c \sum_{l=P}^{\infty} \mathbb{P}\{P_n = l\} / \sqrt{l}. \end{aligned}$$

Since $|F_A(x) - F_Y(x)|$ is given by

$$\left| \int_{-\infty}^{2\pi x/h} (f_X \star (g_1 - f_1) \star g_2 \star \dots \star g_N + \dots + f_X \star f_1 \star f_2 \star \dots \star (g_N - f_N))(y) dy \right|,$$

it follows that

$$|F_A(x) - F_Y(x)| \leq \sum_{n=0}^N \|G_n - F_n\|_{\infty} \leq c \sum_{n=0}^N \sum_{l=P}^{\infty} \mathbb{P}\{P_n = l\} / \sqrt{l} \leq c \frac{N+1}{\sqrt{P}},$$

where we have used the following inequality for the convolution product

$$\left| \int_{-\infty}^z (f \star g)(y) dy \right| \leq \|G\|_{\infty}$$

for any probability density functions f and g , where $G(x) = \int_{-\infty}^x g(y) dy$. \square

The Logistic Normal approximation simulation method we propose is the random variable $Y_N(h)$ defined in Theorem 3.1. We can further augment this method by also simulating the tail, i.e. by simulating $Y_N(h) + (a/\sqrt{3 \cdot 2^{N+1}})(h/2)Z$, where $Z \sim N(0, 1)$. Theorem 3.1 demonstrates that the error in the distribution function corresponding to $A(h)$ as a result of making the Normal replacement $Y_N(h)$, has an upper bound that scales like $(N+1)/\sqrt{P}$. In practice, as we see from our simulations in Figure 2, this upper bound is somewhat pessimistic. On the other hand, if we were to be overly optimistic and supposed that this Normal replacement (of the large sum of Logistic random variables) did not impact the accuracy of the simulation, then we would conclude that above a certain computational effort threshold, the complexity of the Logistic Normal approximation simulation method would be $\log \varepsilon$ —we only need one uniform random variable for each Normal sample and the terms at each order scale like $1/2^n$. In reality the complexity lies somewhere between ε^{-1} (or $\varepsilon^{-1/2}$ when we include the tail approximation) and $\log \varepsilon$.

4. Exponential product approximation

A Logistic random variable can be decomposed as the difference of the logarithms of two independent Exponential random variables. We exploit this result to derive a highly effective sampling method for a large sum of logistic random variables.

Indeed, the sum of P independent standard Logistic random variables can be represented by

$$\sum_{k=1}^P \text{Logistic}_k(1) \sim \log \left(\prod_{i=1}^P \text{Exp}_i(1) \right) - \log \left(\prod_{j=1}^P \text{Exp}_j(1) \right).$$

Hence we are left with the question of whether we can efficiently sample from the probability density function associated with the logarithm of the product of P standard Exponential random variables (with unit means). In fact the density function for the logarithm of a product of P Exponential random variables can be computed via the the inverse Mellin transform as follows (see for example Nadarajah 2009; p. 654). We briefly recall some basic results on products of random variables and the Mellin transform. If two random variables X and Y have densities f_X and f_Y , respectively, then their product XY has density given by the *product convolution*

$$\int_{\mathbb{R}} f_X \left(\frac{x}{y} \right) f_Y(y) \frac{1}{y} dy.$$

The Mellin transform of a function has many properties analogous to that for the Fourier and Laplace transforms—see Flajolet, Gourdon and Dumas (1995) for more details. Importantly the Mellin transform of the product convolution of two functions is given by the real product of their Mellin transforms. Hence since the density of the product of P Exponential random variables is the P -fold product convolution of the corresponding densities, its Mellin transform is the P -fold real product of the Mellin transform of the standard Exponential random variable density which is the Gamma function. The probability density function $\varphi = \varphi(x)$ corresponding to product of P standard Exponential random variables is thus given by the inverse Mellin transform of the P th power of the Gamma function defined for all $x > 0$ by

$$\varphi(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} x^{-z} (\Gamma(z))^P dz,$$

where $c > 0$ is any real constant. The probability density function $\phi_P = \phi_P(x)$ corresponding to the *logarithm* of such a product of Exponential random variables is given by $\phi_P(x) = \varphi(e^x) e^x$. In other words we have

$$\phi_P(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{x(1-z)} (\Gamma(z))^P dz.$$

We can explicitly compute the contour integral above using the theory of residues from complex analysis to derive a series representation for the density ϕ_P ; see Appendix A. Unfortunately the coefficients of this series grow so fast with P , that its radius of convergence and range of practical evaluation makes it redundant for values of P greater than 10. We can also derive the leading order asymptotic form for ϕ_P for large x . However again unfortunately, for large P , the range of large values of x for which this asymptotic form is useful, becomes severely restricted. In practical terms, the most useful representation is the leading order asymptotic form for ϕ_P for large P , uniform in x , derived using the method of steepest descents.

Theorem 4.1 (Logarithm of Product Exponentials density). *For large P , uniformly in x , the probability density function $\phi_P = \phi_P(x)$ for logarithm of the product of $P \in \mathbb{N}$ Exponential random variables, has the leading order asymptotic form*

$$\phi_P(x) \sim \frac{(\Gamma(\psi^{-1}(x/P)))^P \cdot e^{x(1-\psi^{-1}(x/P))}}{\sqrt{2\pi P \psi'(\psi^{-1}(x/P))}},$$

where $\psi = \Gamma'/\Gamma$ is the Digamma function. The next order term is asymptotically smaller than a factor $1/P$ times a functional form similar to the numerator for ϕ_P above (rather than the factor $1/\sqrt{P}$ shown for ϕ_P).

Proof. To determine the asymptotic behaviour for ϕ_P for large P , we apply the method of steepest descents. Recall that $\phi_P = \phi_P(x)$ is given for any real $c > 0$ by

$$\phi_P(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{x(1-z)} (\Gamma(z))^P dz = \frac{e^x}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{P((-x/P)z + \log \Gamma(z))} dz.$$

Let $h(z) := (-x/P)z + \log \Gamma(z)$ denote the exponent function in the integrand. Note that we retain the ratio $(-x/P)$ as we wish to include the possibility that $(-x/P)$ might be order one or even asymptotically large. Since P and x are real, we note that $h = h(z)$ has a unique saddle point given by $z = z_* := \psi^{-1}(x/P) > 0$ where $\psi = \Gamma'/\Gamma$ is the Digamma function. Since the argument in the above contour integral is analytic in the right-half complex plane, we can deform the contour of integration to one of constant phase that passes through $z = z_*$ on the real axis. Then we apply the usual Laplace method arguments. We shrink the range of the contour of integration to a small interval around $z = z_*$, replace $h(z)$ by its quadratic Taylor expansion, then extend the range of integration out again and utilize standard area estimates for Gaussian functions. This standard procedure reveals that

$$\phi_P(x) \sim \frac{e^x}{2\pi i} \cdot i \cdot \sqrt{\frac{-2\pi}{Ph''(z_*)}} \cdot e^{Ph(z_*)} + \mathcal{O}\left(\frac{e^{Ph(z_*)}}{P}\right),$$

where $h''(z_*)$ denotes the second derivative of h along the constant phase contour (orthogonal to the real axis) at $z = z_*$. Substituting the form for $h = h(z)$ and that $z_* = \psi^{-1}(x/P)$ gives the large P result of the theorem. \square

As with the Logistic Normal approximation, the idea here is to more efficiently sample the large sums of Logistic random variables, on average $\mathbb{E}\{P_n\} = \frac{1}{2}a^22^n$, at each order; especially at higher orders. Hence when the Poisson sample $P_n \geq P$ at order n , for some predetermined $P \gg 1$, we will instead draw samples from the distribution function Φ_{P_n} for the logarithm of P_n standard Exponential random variables—corresponding to the density function ϕ_{P_n} . The only practical form for ϕ_{P_n} for sampling when $P_n > 10$ is the asymptotic form given in Theorem 4.1. However, unfortunately an analytic form for distribution function Φ_{P_n} is not available. To overcome this we proceed as follows. We tabulate the corresponding inverse distribution function $\Phi_{10^\ell}^{-1}$ for $\ell = 2, 3, \dots$. This needs to be done only once and as accurately as we possibly can. Then given any $P_n \sim \text{Poisson}(\frac{1}{2}a^22^n)$ such that $P_n \geq P$, with say $P = 10^2$, we can decompose $P_n = a_1 + a_2 \cdot 10^2 + a_3 \cdot 10^3 + \dots$. Here

a_1, a_2, \dots are simply the decimal digits of P_n . In the Exponential product approximation method we propose then, for each n , when $P_n \geq P$ we sample a_1 Logistic random variables X_1, \dots, X_{a_1} directly. Then using interpolation with the table for $\Phi_{10^2}^{-1}$ we sample a_2 pairs of random variables $U_{2,1}, \dots, U_{2,a_2}$ and $V_{2,1}, \dots, V_{2,a_2}$ from Φ_{10^2} . We further sample a_3 pairs of random variables $U_{3,1}, \dots, U_{3,a_3}$ and $V_{3,1}, \dots, V_{3,a_3}$ using interpolation from our table for $\Phi_{10^3}^{-1}$, and so forth. Hence for $P_n \geq P$, we replace the Logistic random variable accumulator as follows

$$\sum_{k=1}^{P_n} X_{n,k} \approx \sum_{j=1}^{a_1} X_j + \sum_{\ell \geq 2} \sum_{j=1}^{a_\ell} U_{\ell,j} - V_{\ell,j}.$$

In practice we restricted $\ell \leq 5$. The only error in this replacement is in the asymptotic approximation of the cumulative distribution function, the accuracy of the tables and the interpolation. Above the threshold P , the number of uniform random we require at each order is two times $a_1 + a_2 + a_3 + \dots$. Hence, since at each order the Logistic accumulator is scaled by the factor 2^n , the complexity of our Exponential product approximation method is in principle $\log \varepsilon$. We can of course augment this method by simulating the tail sum as before.

5. Simulation details

We provide here details of our implementation underlying Figure 2 and at the same time explain more carefully what we observe therein. Our Lévy and basic Logistic simulation are straightforward truncations of the corresponding representations given in the introduction and Theorem 2.1. Note that in all cases we performed $L = 5 \times 10^8$ simulations—as large as was practically possible. Monte–Carlo simulation variance is of order $1/L$. From our experience and as can be seen in Figure 2, there is significant Monte–Carlo simulation noise roughly below 10^{-4} , which we should take into account when interpreting the mean-square errors therein. Note that we measure the error as the absolute value of the sample variance minus the true variance $(1 + a^2)h^2/12$, where with random Wiener increments $\Delta W^1(h)$ and $\Delta W^2(h)$ for each sample, we set $a^2 = \mathbb{E}\{(\Delta W^1(h))^2 + (\Delta W^2(h))^2\} = 2h$. We see in Figure 2 that the Lévy method provides Lévy area samples for low accuracies and low computational effort. The slope of the error versus the actual computational effort for the Lévy method is -1 as it should be (the variance of the tail asymptotically decays like $1/n$). If we also simulate the tail for the Lévy method, then the mean-square error should decay like $1/n^2$; see Kloeden, Platen & Wright (1992) or Wiktorsson (2001). Note that we do not directly observe this improved accuracy/convergence in Figure 2 as we compute the sample variance for the Lévy method with independent tail approximations. Thus we expect to see the error of the Lévy method with tail approximation to be similar to the error of the Lévy method itself, shifted by the variance of the tail approximation—an independent computation confirms this. Indeed such a shift is observed for all the methods we implemented when we also simulate the tail, except that the shift is much more dramatic for all the Logistic expansion methods. In the case of the basic Logistic method, we can only generate Lévy area samples if we’re willing to invest computational effort above a (in principle quantifiable) threshold. In the basic Logistic approximation, the order scales exponentially—the terms at each order scale like

2^{-n} but we must invest on average a factor $\mathbb{E}\{P_n\} = \frac{1}{2}a^22^n$ amount of effort at each order as well. In Figure 2, low order basic Logistic approximations scale favourably as it's relatively cheap to sum a medium-sized number $\frac{1}{2}a^22^n$ of Logistic random variables (whilst the terms scale like 2^{-n} at each order). However eventually, as we can see in Figure 2, we reach the asymptotic behaviour we expect with a slope of -1 . When we also simulate the tail, we get a dramatic improvement in accuracy, so much so that the results lie in the regime where we know there is significant Monte–Carlo simulation noise. Note that we did not include in Figure 2 simulations of the Rydén–Wiktorsson method. The computational effort required to achieve the corresponding accuracies we have indicated in Figure 2 meant that the Rydén–Wiktorsson method, even including simulating the tail, was not competitive. Note also that for all our Logistic expansion based methods, we used the standard algorithm given in Knuth (1998) to simulate the Poisson random variable P_n when the mean $\frac{1}{2}a^22^n \leq 10^2$ and the Normal approximation otherwise.

In Figure 2 we also indicate the performance of our Normal approximation and Exponential product approximation methods. In the case of the Normal approximation, at each order, when the Poisson sample $P_n > 10^2$ we invoked the Central Limit Theorem and replaced the Logistic sum by the suitably matched Normal random variable as indicated in §3. This effect is strikingly indicated by the sharp downturn in the Normal approximation plot in Figure 2. We also observe that the accuracy of the method is robustly uncompromised by this Normal replacement. Note that including a tail approximation generates a much more accurate sampling method. As we indicated above, an independent computation demonstrates that the error shown in Figure 2 for the Normal approximation with tail is the same as the Normal approximation shifted by the variance of the tail approximation. Note that the Normal approximation with tail carries us into the regime where the Monte–Carlo simulation noise is significant.

To implement the Exponential product approximation some preparation is required. We constructed tables for the inverse distribution functions Φ_P^{-1} for the logarithm of the product of P Exponential random variables for $P = 10^2, 10^3, 10^4$ and 10^5 as follows. Using the large P asymptotic approximation for ϕ_P we derived in Theorem 4.1, we constructed Φ_P by integrating from the left and right ends, respectively, as follows (we justify this presently):

$$\Phi_P(x) = \int_{-\infty}^x \phi_P(y) dy \quad \text{and} \quad \Phi_P(x) = 1 - \int_x^{\infty} \phi_P(y) dy.$$

We used the trapezium rule with very small stepsizes to approximate these two expressions and the results are given in Figure 3. In the left set of panels of the Figure we plot the distribution functions Φ_P constructed in this way. The blue curve corresponds to integrating from the left and the green to integrating from the right. An empirically computed cumulative distribution function is shown in red. The accuracy of the integrated asymptotic form is immediately apparent; as is the reason for independently integrating from either end to minimize the accumulated quadrature error. For the case $P = 10^2$, the magenta curve shown corresponds to constructing the distribution function by a quadrature approximation of the contour integral form for $\Phi_P = \Phi_P(x)$ for each x . For small x this approximation breaks down and for larger values of P it becomes impractical. In the right set of panels in Figure 3 we show the difference of the left and right integral approximations

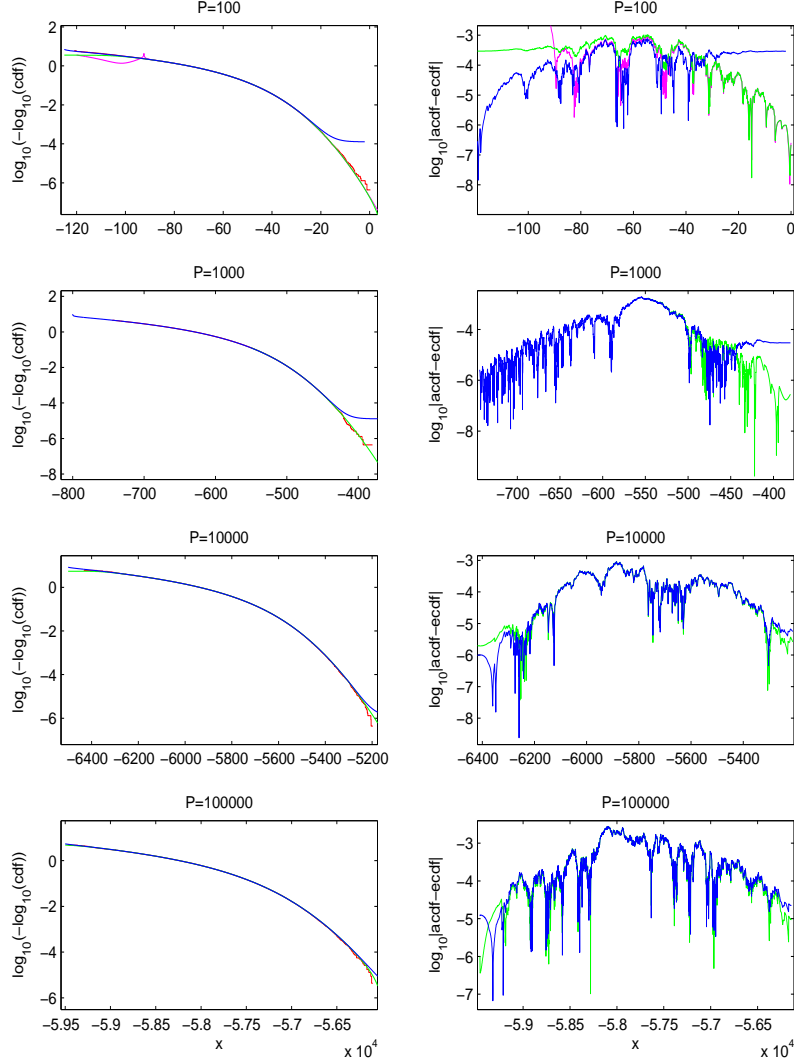


Figure 3. The panels on the left show the asymptotic approximations (acdf: from the left, blue, from the right, green) and the empirical cumulative distribution function (ecdf: red; for which we used a ‘stairs’ plot) for the logarithm of the product of P standard Exponential random variables. Note we have used a log-of-minus-log scale. The panels on the right give an indication of the error of the the asymptotic approximations compared to the empirical cumulative distribution function (with the same colour attribution). The magenta curve in the top two panels represents an approximation to the cumulative distribution function constructed by directly numerically computing the contour integral for ϕ_P .

to the empirically computed cumulative distribution function which indicate errors less than 10^{-3} . We spliced together the left and right approximations, numerically inverted the result, and then used linear interpolation to construct the inverse $\Phi_P^{-1} = \Phi_P^{-1}(x)$ cumulative distribution functions using $10^6 + 1$ points uniformly spaced on $[0, 1]$. In other words, the tables each contain $10^6 + 1$ values for Φ_P^{-1} at $x_m = m/10^6$ for $m = 0, 1, 2, \dots, 10^6$. In the tables corresponding to $P = 10^3$ and above, we crudely set $\Phi_P^{-1}(1) = 1$ whilst we set $\Phi_{10^2}^{-1}(1) = 10$. Given the range of values of Φ_P^{-1} as $x \rightarrow 1$ we can infer from the left set of panels in Figure 3 this replacement of $+\infty$ seems reasonable and in practice from what we observe in Figure 2 this did not affect accuracy. The tables thus constructed in this way, allow us to rapidly sample from Φ_P for any of the values of P we've indicated, as follows. We generate a uniform random variable U . The integer part of $U \times 10^6$ immediately indicates the position in the table of interest to us. We use linear interpolation between that position and the next to generate a sample from Φ_P . Thus, with these tables and linear interpolation method for sampling from them, we implemented the Exponential product approximation as indicated at the end of §4. We can of course also include a tail approximation. The results shown in Figure 2 are just as dramatic as for the Normal approximation, making both these methods with tail simulation the ones of choice for high accuracy simulations.

To conclude we remark that the improved sampling methods for the Lévy area we have introduced have important implications for our theoretical work on the algebraic structure of stochastic differential equations and concomitant universally accurate integrators (see Malham & Wiese 2009 and Ebrahimi-Fard *et. al.* 2011) as well as for more practical work on the pricing of financial derivatives (see Malham & Wiese 2011).

Appendix A. Series representation

Applying the theory residues to the contour integral form for the probability density function ϕ_P , corresponding to the logarithm of the product of P standard Exponential random variables, we can explicitly derive a series representation for it. For convenience, we introduce the *discrete convolution product*. For two sequences $a = \{a_0, a_1, a_2, \dots\}$ and $b = \{b_0, b_1, b_2, \dots\}$ this is the sequence whose n th term is given by $(a * b)_n := a_n b_0 + a_{n-1} b_1 + \dots + a_0 b_n$, for $n = 0, 1, 2, \dots$. If a and b are the coefficients of the corresponding powers in a power series, say $a_0 + a_1 x + a_2 x^2 + \dots$ and $b_0 + b_1 x + b_2 x^2 + \dots$, then the coefficients of the product of these two power series are $a * b$.

Theorem 5.1 (Series representation). *For the probability density function $\phi_P = \phi_P(x)$ for the logarithm of the product of $P \in \mathbb{N}$ Exponential random variables we have the following explicit series representation*

$$\phi_P(x) = \sum_{k=0}^{P-1} x^k \sum_{n \geq 0} a_k(n) e^{(n+1)x},$$

where

$$a_k(n) := \frac{(-1)^{nP+k}}{(n!)^P k!} \cdot (q * r(n))_{P-1-k}.$$

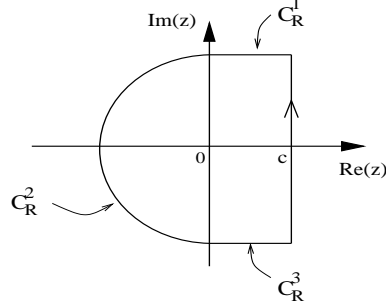


Figure 4. The contour \mathcal{C}_R with sections \mathcal{C}_R^1 , \mathcal{C}_R^2 and \mathcal{C}_R^3 used in the proof of Theorem 4.1.

Here if $\Gamma^{(k)}(1)$ is the k th derivative of the Gamma function evaluated at one, and if C_k^P denotes the usual Binomial coefficient, P choose k , then

$$q := \left(\{ \Gamma^{(k)}(1)/k! \}_{k \geq 0} \right)^{*P}$$

and

$$r(n) := \{C_k^P\}_{k \geq 0} * \{C_k^P/2^k\}_{k \geq 0} * \cdots * \{C_k^P/n^k\}_{k \geq 0}.$$

Proof. To derive the power series representation for the probability density function ϕ_P , we compute the inverse Mellin transform φ of Γ^P . Here $\varphi = \varphi(x)$ corresponds to the probability density function of the product of P standard Exponential random variables. At the very end we use that $\phi_P(x) = \varphi(e^x) e^x$.

To analytically compute the inverse Mellin transform of Γ^P we use the theorem of residues for the contour $\mathcal{C}_\infty = \lim_{R \rightarrow \infty} \mathcal{C}_R$ shown in Figure 4; we found Gajjar (2010, p. 111–135) very useful here. Note that we have

$$\begin{aligned} \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} x^{-z} (\Gamma(z))^P dz &= \frac{1}{2\pi i} \int_{\mathcal{C}_\infty} x^{-z} (\Gamma(z))^P dz \\ &= \sum_{n \geq 0} \text{res} \left(x^{-z} (\Gamma(z))^P : z = -n \right). \end{aligned}$$

Here we have used that the contributions to the contour integral from the components of the contour \mathcal{C}_R^1 , \mathcal{C}_R^2 and \mathcal{C}_R^3 are zero. Our goal now is to determine the residue of $x^{-z} (\Gamma(z))^P$ at each of its poles $z = -n$ for $n = 0, 1, 2, \dots$. Note these are poles of order P due to $\Gamma(z)$ have simple poles at these points. Our goal now is thus to derive the Laurent series in powers of $(z+n)$ for $x^{-z} (\Gamma(z))^P$ centred at each of these poles, and to determine the coefficients of $(z+n)^{-1}$ corresponding to the residues. For convenience we set $\zeta := z+n$. The Taylor series for x^{-z} centred at $z = -n$ is given by

$$x^{-z} = x^n \cdot \exp(-\zeta \log x) = x^n \sum_{k \geq 0} \frac{(-\log x)^k}{k!} \zeta^k = x^n \sum_{k \geq 0} \ell_k(x) \zeta^k,$$

where we set $\ell_k(x) = (-\log x)^k/k!$ for all $k \geq 0$. Then using the Taylor series for $\Gamma(1+z)$ centred at $z=0$, we have

$$\begin{aligned}\Gamma(z) &= z^{-1}(z+1)^{-1} \cdots (z+n)^{-1} \Gamma(1+n+z) \\ &= (\zeta-n)^{-1}(\zeta-n+1)^{-1} \cdots \zeta^{-1} \cdot \sum_{k \geq 0} \frac{\Gamma^{(k)}(1)}{k!} \zeta^k.\end{aligned}$$

Taking the P power of this expression, and using the definition for the coefficients $\{q_k\}$ and $\{r_k(n)\}$ given in the theorem, we get

$$\begin{aligned}(\Gamma(z))^P &= (\zeta-n)^{-P}(\zeta-(n-1))^{-P} \cdots (\zeta-1)^{-P} \zeta^{-P} \cdot \left(\sum_{k \geq 0} \frac{\Gamma^{(k)}(1)}{k!} \zeta^k \right)^P \\ &= \frac{(-1)^{Pn}}{(n!)^P} \cdot (1-\zeta/n)^{-P} (1-\zeta/(n-1))^{-P} \cdots (1-\zeta)^{-P} \zeta^{-P} \cdot \sum_{k \geq 0} q_k \zeta^k \\ &= \frac{(-1)^{Pn}}{(n!)^P} \cdot \zeta^{-P} \cdot \left(\sum_{k \geq 0} r_k(n) \zeta^k \right) \left(\sum_{k \geq 0} q_k \zeta^k \right) \\ &= \frac{(-1)^{Pn}}{(n!)^P} \cdot \zeta^{-P} \cdot \sum_{k \geq 0} (r(n) * q)_k \zeta^k.\end{aligned}$$

Thus we are now in a position to compute the Laurent series for $x^{-z}(\Gamma(z))^P$ centred at $\zeta=0$ corresponding to $z=-n$. Directly computing, we have

$$\begin{aligned}x^{-z}(\Gamma(z))^P &= \frac{(-1)^{Pn}}{(n!)^P} \cdot \zeta^{-P} \cdot x^n \cdot \left(\sum_{k \geq 0} \ell_k(x) \zeta^k \right) \left(\sum_{k \geq 0} (r(n) * q)_k \zeta^k \right) \\ &= \frac{(-1)^{Pn}}{(n!)^P} \cdot x^n \cdot \sum_{k \geq 0} \left(\ell(x) * (r(n) * q) \right)_k \cdot \zeta^{k-P},\end{aligned}$$

where $\ell(x) = \{\ell_k(x)\}_{k \geq 0}$. The residue of $x^{-z}(\Gamma(z))^P$ at $\zeta=0$ is generated by the term $k=P-1$, i.e. we have

$$\text{res}\left(x^{-z}(\Gamma(z))^P : z=-n\right) = \frac{(-1)^{Pn}}{(n!)^P} \cdot x^n \cdot (\ell(x) * r(n) * q)_{P-1}.$$

Using the explicit sum for the first discrete convolution on the right-hand side, summing over all the residues $z=0, -1, -2, \dots$, and using that $\phi_P(x) = \varphi(e^x)e^x$, generates the form for the density function ϕ_P given in the theorem. Note that $r(0) = \{1, 0, 0, \dots\}$ and so $(\ell(x) * r(0) * q)_{P-1} \equiv \ell_0(x)q_{P-1} + \ell_1(x)q_{P-2} + \cdots + \ell_{P-1}(x)q_0$. \square

This explicit series representation is useful and accurate for relatively small values of P . In particular ϕ_2 matches the known modified Bessel function form $2K_0(2e^{x/2})e^x$ for the probability density function when $P=2$. Further, integration by parts soon generates an analytic series expansion for the corresponding distribution function. Unfortunately however, the coefficients q grow factorially with P making the series impractical for P greater than 10.

As mentioned in the main text, we can also derive the leading order asymptotic form for ϕ_P for large x . Consider the product convolution form for density of the product of P standard Exponential random variables given by

$$\varphi(x) = \int_0^\infty \cdots \int_0^\infty \exp\left(-\frac{x}{y_2 \cdots y_P}\right) \frac{\exp(-(y_2 + \cdots + y_P))}{y_2 \cdots y_P} dy_2 \cdots dy_P.$$

After using the substitution $y_i = u_i x^{1/P}$, we can then apply Laplace's method. Noting that the exponent function has a global maximum at $y_2 = \cdots = y_P = 1$ in the range of integration, then up to exponentially small errors, we have that

$$\phi_E(x) \sim \frac{\exp(-Px^{1/P})}{x^{(P-1)/2P}} \cdot \int_{-\infty}^\infty \cdots \int_{-\infty}^\infty \exp\left(-\sum_{i=1}^{P-1} \tau_i^2 - \sum_{i \neq j=1}^{P-1} \frac{1}{2} \tau_i \tau_j\right) d\tau_1 \cdots d\tau_{P-1}.$$

The matrix generating the quadratic form in the exponent in the integrand has one eigenvalue equal to $P-1$ with the remaining eigenvalues all equal to one. A linear change of coordinates corresponding to the orthonormal similarity transformation of the matrix generating the quadratic form, and using that $\phi_P(x) = \varphi(e^x) e^x$ generates the asymptotic form

$$\phi_P(x) \sim \frac{(2\pi)^{\frac{P-1}{2}}}{\sqrt{P}} \cdot \exp(-P \exp(x/P) + (P+1)x/2P).$$

For large P though, the range of large values of x for which this approximation is useful, is severely restricted.

Lastly, we remark that an analytical form for the probability density function corresponding to a sum of Logistic random variables is given in George and Mudholkar (1983). We have not pursued this approach, though one could in principle investigate whether a computationally feasible method to sample a sum of Logistic random variables could be derived from this form.

References

- Cruzeiro, A.B., Malliavin, P. and Thalmaier, A. (2004). Geometrization of Monte-Carlo numerical analysis of an elliptic operator: strong approximation. *C.R. Acad. Sci. Paris Ser. I* **338** 481–486.
- Ebrahimi-Fard, K., Lundervold, A., Malham, S.J.A., Munthe-Kaas, H.Z. and Wiese, A. (2011). Algebraic structure of stochastic simulation. Preprint.
- Flajolet, P., Gourdon, X. and Dumas, P. (1995). Mellin transforms and asymptotics: Harmonic sums. *Theoretical Computer Science* **144**(1-2) 3–58.
- Gaines, J. and Lyons, T.J. (1994). Random generation of stochastic area integrals. *SIAM J. Appl. Math.* **54**(4) 1132–1146.
- Gaines, J. and Lyons, T.J. (1997). Variable step size control in the numerical solution of stochastic differential equations. *SIAM J. Appl. Math.* **57**(5) 1455–1484.
- Gajjar, J. (2010). Mathematical Methods. MAGIC lecture notes. These are available from <http://maths.dept.shef.ac.uk/magic/course.php?id=145>
- George, E.O. and Mudholkar, G.S. (1983). On the convolution of logistic random variables. *Metrika* **30** 1–13.
- Gillsing, H. and Shardlow, T. (2007). SDELab: A package for solving stochastic differential equations in MATLAB. *J. Comput. Appl. Math.* **205** 1002–1018.

- Karhunen, K. (1947). Über lineare Methoden in der Wahrscheinlichkeitsrechnung. *Ann. Acad. Sci. Fenn., Ser. A, I, Math. Phys.* **37** 3–79.
- Kloeden, P.E., Platen, E. and Wright, W. (1992). The approximation of multiple stochastic integrals. *Stochastic Anal. Appl.* **10** 431–441.
- Knuth, D.E. (1998). *The art of computer programming, vol 2: Seminumerical algorithms*, Addison-Wesley, Reading, Mass., Third edition.
- Loève, M. (1945). Sur les fonctions aléatoires stationnaires du second ordre. *Revue Scientifique* **83** 297–310.
- Levin, D. and Wildon, M. (2008). A combinatorial method for calculating the moments of the Lévy area. *Trans. AMS* **360**(12) 6695–6709.
- Levy, P. (1951). Wiener's random function and other Laplacian random functions. *Second Symposium of Berkeley. Probability and Statistics*, UC Press, 171–186.
- Lord, G., Malham, S.J.A. & Wiese, A. 2008 Efficient strong integrators for linear stochastic systems, *SIAM J. Numer. Anal.* **46**(6), 2892–2919
- Malham, S.J.A. and Wiese, A. (2008). Stochastic Lie group integrators. *SIAM J. Sci. Comput.* **30**(2) 597–617.
- Malham, S.J.A. and Wiese, A. (2009). Stochastic expansions and Hopf algebras. *Proc. R. Soc. A* **465** 3729–3749.
- Malham, S.J.A. and Wiese, A. (2011). Positive volatility simulation in the Heston model, submitted. (arXiv:0802.4411v4)
- Milstein, G. N. (1994). *Numerical integration of stochastic differential equations*, Mathematics and its applications, Kluwer Academic Publishers.
- Nadarajah, S. (2009). Exact distribution of the product of two or more logistic random variables. *Methodol. Comput. Appl. Probab.* **2009** **11** 651–660.
- Pitman, J. and Yor, M. (2003). Infinitely divisible laws associated with hyperbolic functions. *Canad. J. Math.* **55**(2) 292–330.
- Rydén, T. and Wiktorsson, M. (2001). On the simulation of iterated Itô integrals. *Stochastic Process. Appl.* **91** 151–168.
- Stump, D. M. and Hill, J. M. (2005). On an infinite integral arising in the numerical integration of stochastic differential equations. *Proc. R. Soc. A* **461** 397–413.
- Wiktorsson, M. (2001). Joint characteristic function and simultaneous simulation of iterated Itô integrals for multiple independent Brownian motions. *Ann. Appl. Probab.* **11**(2) 470–487.